

**Kempeneer/van Dooren Using numbers that do not count: how the latent functions of performance indicators explain their success**

International Review of Administrative Sciences  
Band 87 (2), Juni 2021, S. 364-379

**Der Gebrauch der Zahlen, die nicht zählen: wie die latenten Funktionen der Leistungsindikatoren ihren Erfolg erklären**

Seit der Weltfinanzkrise sind Stresstests wesentlicher Bestandteil des Arsenal der Regulierungsbehörden zur Kontrolle und Aufrechterhaltung der globalen finanziellen Stabilität geworden. In Europa führt die Europäische Bankaufsichtsbehörde (EBA) in enger Zusammenarbeit mit der Europäischen Zentralbank (EZB) systemweite Stresstests durch, um die Fähigkeit der großen, sog. systemischen Banken zu testen, mit negativen Szenarien fertig zu werden, was es ermöglicht, Banken von Schlechtleistern bis zu Best Performern einzuordnen.

Auf den ersten Blick ist der Erfolg der Stresstests leicht erklärbar. In dem Maße, wie das Bewusstsein für Risiken steigt, wächst auch die Nachfrage nach Instrumenten, die diese Risiken bemessen und kontrollieren. Der Erfolg der Leistungsindikatoren beruht auf ihrer Fähigkeit, komplexe Systeme zu vereinfachen, zu standardisieren, zu vergleichen und zu kontrollieren. Der EU-weite Stresstest antwortet auf diesen Kontrollbedarf. Die Stresstests sollen auf einen Blick die Banken anzeigen, die einer Krise widerstehen würden.

Allerdings haben jahrzehntelange Forschungen auch die nachteiligen Auswirkungen von Leistungsindikatoren gezeigt: sie seien ungenau, verführten zu Glücksspiel, demotivierten die Mitarbeiter, verzerrten Messwerte und vermochten nicht, die Performance substantziell zu verbessern. Forschungen haben endlich auch den Missbrauch von Leistungsdaten als Schlüsselfaktor identifiziert, der die Leistungsverbesserung behindert. Einige Forscher gehen so weit zu behaupten, dass evidenzbasierte und rationale Politikgestaltung ein Mythos ist, weil politische Arbeit eben grundlegend „politisch“ sei. Nachteilige Wirkungen zeigten sich vor allem, wenn Leistungsindikatoren von hoher Anreizwirkung benutzt würden. Diese Kritik hat die Stresstests, vor allem seitens der Finanzpresse, nicht verschont. Stresstests hätten ökonomisch keinen Sinn, seien gegenüber bestimmten Banken voreingenommen und wenig mehr als Kommunikationsübungen zur Beruhigung der Finanzmärkte. Natürlich gibt es Studien, die zeigen, wie Performancemessungen Leistungsergebnisse positiv beeinflussen. Aber es

mangelt an Verständnis, warum in einigen Fällen Performance-Indikatoren die Leistungsergebnisse verbessern und in anderen nicht.

Trotz dieser verschiedenen und offensichtlich widersprüchlichen Effekte von Leistungsindikatoren propagieren Regulierungsbehörden sie als unerlässliche Werkzeuge für die Regulierung. Im Lauf der letzten Jahrzehnte hat die Nutzung von Performance-Indikatoren in der globalen Ordnungspolitik sich stark ausgebreitet.

Der vorliegende Beitrag soll den fortwährenden Erfolg von Leistungsindikatoren erklären und erforschen, unter welchen Bedingungen sie trotz ihrer Schwächen und Fehlfunktionen Leistungsergebnisse verbessern können.

Es ist festzustellen, dass Leistungsindikatoren wichtige – verborgene (latente) – Funktionen haben, die bisher in der Literatur kaum behandelt wurden. Dabei halten wir uns an die funktionale Soziologie von Merton (1968). Merton unterscheidet zwischen offensichtlichen Funktionen (beabsichtigte und positive Effekte), Fehlfunktionen (unbeabsichtigte und negative Effekte) und latenten (unbeabsichtigten positiven Effekten) Funktionen. Die Performance-Literatur befasst sich typischerweise mit den offensichtlichen Funktionen und den Fehlfunktionen in Bezug auf die Performance-Informationen.

Die hier vorliegende Fallstudie untersucht die latenten Funktionen von Stresstests. Der Prozess der Auswertung des Stresstests bewirkt eine latente Verbesserung der Leistungsergebnisse. Um den Stresstest zeitnah zu vollenden, haben die Banken ihre internen Risikomanagementsysteme professionalisiert, indem sie in verbesserte Datenqualität und IT-Systeme sowie eine bessere Koordination zwischen den verschiedenen Risikogebieten investiert haben. Damit haben sie unbeabsichtigt ihre Performanceergebnisse verbessert. Des Weiteren wurde herausgefunden, dass die latenten Funktionen viel wichtiger für die Performanceergebnisse sind als oft angenommen. Schließlich wurde festgestellt, dass bestimmte Fehlfunktionen, wie Ungenauigkeit der Ergebnisse für die Performanceresultate nicht notwendigerweise schädlich sind.

## Nachweisbare Funktionen von Performance-Indikatoren

In der öffentlichen Regierungs- und Verwaltungsführung wimmelt es von Indikatoren. Die NPM-Reformen haben in allen Ecken der Verwaltung zu einer Verbreitung von Indikatoren geführt. Das NPM umfasst eine Unmenge von Prinzipien, die alle auf eine Verbesserung der Performance im öffentlichen Sektor abzielen, indem sie ihn effizienter und zielorientierter machen. Das geschieht durch Entflechtung, Wettbewerb und Anreize. Damit leistungsorientierte Regulierung funktioniert, muss die Performance notwendig gemessen werden, und hier betreten die Leistungsindikatoren die Szene. Viele der im Rahmen des

NPM herausgestellten Anreize sind nur anwendbar, wenn es quantitative Indikatoren gibt.

Leistungskennziffern finden in die Werkzeugkästen der Regulierer leicht ihren Weg, wenn sie auf einen Blick Informationen liefern. Sie bieten objektive Bewertungskriterien, welche Organisationen ihre Ziele erreichen und welche unter ihren Möglichkeiten bleiben. Diese Informationen können benutzt werden, um Ergebnisse zu verbessern, indem sie Wettbewerb zwischen Organisationen fördern, Ressourcen entsprechend der Performance zuteilen und die Rechenschaftspflicht ausweiten. Darüber hinaus ermöglicht die Einfachheit der Performance-Indikatoren eine knappe Kommunikation innerhalb und außerhalb der Verwaltung und damit, jedenfalls theoretisch, mehr Transparenz und Rechenschaft.

### Fehlfunktionen der Performance-Indikatoren

Trotz edler Absichten und hochfliegender Hoffnungen gibt es eine zunehmende Zahl von kritischen Stimmen zu den dysfunktionalen Wirkungen des NPM und der Performance-Indikatoren, die sich für eine NPM-Reform aussprechen. Nach mehr als drei Jahrzehnten der Leistungsbemessung in der öffentlichen Verwaltung erwähnen selbst wohlwollende Analytiker die abträglichen Effekte der NPM-Reformen, vor allem in Bezug auf Leistungsindikatoren. Obwohl in einigen (speziell Entwicklungs-)Ländern NPM-Reformen noch angewendet werden, haben die meist entwickelten Länder erkannt, dass NPM öffentliche Organisationen nicht effizienter und erfolgreicher gemacht hat. Stattdessen haben NPM-Themen wie Entflechtung, Wettbewerb und Anreizsysteme zu in Schubladendenken verhafteten öffentlichen Einrichtungen geführt, das kollektives Handeln verhindert, zu widersinnigen Quasi-Markt-Mechanismen und einer Fixierung auf mittelfristige Organisationsziele geführt, die Dienstleistungen und Effektivität überschatten. Auch die Performance-Indikatoren wurden vielfach kritisiert.

Eine erste Fehlfunktion besteht darin, dass Performance-Indikatoren zum Tunnelblick führen können. Sie tendieren dazu, sich auf leicht quantifizierbare Performance-Dimensionen zu fokussieren und dadurch den Blickwinkel auf Politik und politische Diskussionen auf kleine und oft nicht repräsentative Aspekte der Politik zu verengen. Dem wird hinzugefügt, dass eine übergroße Abhängigkeit von Größen und Rankings Gefahr läuft, dass die Phänomene, die sie abbilden, hinsichtlich Fortschritt und Richtung bewegliche Ziele sind. Bei den Stresstests sind die leicht quantifizierbaren Risikogebiete wie Kredit- und Marktrisiken stets vorrangig angesprochen worden, während schwieriger quantifizierbare, einzugrenzende und zu definierende Gebiete wie betriebliche Risiken weniger entwickelt sind.

Daneben können Performance-Indikatoren widersinnige Anreize hervorrufen und zu „Glücksspiel“ und Betrug verführen. Es gibt viele Beispiele dafür, wie Daten manipuliert werden. Darüber hinaus sollen Indikatoren auch die Wissbegierde unterdrücken und Lernmöglichkeiten reduzieren. Des Weiteren können Performance-Indikatoren zu einer Zielverschiebung führen, wenn Organisationen sich statt auf ihre Ziele eher auf die sie angeblich bemessenden Indikatoren fokussieren. Performance-Indikatoren vernebeln oft, was am Arbeitsplatz wirklich vor sich geht, und streuen Verdacht und Misstrauen, unterminieren Berufsethik und schaffen eine Reihe von unvorhergesehenen Problemen.

Schließlich werden die durch die Performance-Indikatoren erlangten Informationen häufig nicht für die Entscheidungsfindung benutzt. Häufige Gründe sind die mangelhafte Qualität der Performance-Informationen und das Fehlen wichtiger Daten, aber auch kulturelle und institutionelle Hindernisse. Performance-Indikatoren sind bloß weiterer Papierkram, der sich in Akten und Computer-Ordern verliert.

Das lässt uns mit einem Rätsel zurück: wenn Forschung und Praxis so viele Fehlfunktionen dokumentieren: unter welchen Umständen können Performance-Indikatoren Leistungsergebnisse wirklich verbessern?

## Methodik

Dieser Frage wird nachgegangen durch eine Untersuchung, welchen Beitrag der Stresstest zum Risikoperformance-Management leistet. Es geht um die Frage, wie die Akteure (hier: die Nationalbank von Belgien – NBB –, die Europäische Zentralbank und die Europäische Bankenaufsichtsbehörde- EBA) Stresstests einen Sinn und eine Bedeutung als Steuerungsinstrumente geben. Hierfür wurde der EU-weite Bankenstresstest als hervorragendes Beispiel eines Performance-Indikators, der im Gefolge der Finanzkrise von 2008 in das Zentrum der Risiko-Regulierung geraten ist, ausgewählt. Es wurde auch untersucht, wie sich dieser ziemlich neue Indikator, der jährlichen Veränderungen unterworfen ist, über die Jahre entwickelt hat und wie die verschiedenen Wiederholungsschritte des Stresstests in den Augen der Betroffenen die Performance-Ergebnisse beeinflusst haben.

Es wurden 45 Interviews mit 33 Personen geführt, die im Schnitt 75 Minuten dauerten. Diese verteilten sich auf belgische Banken und Consulting-Firmen, die EZB, die EBA und die NBB.

## Die latenten Funktionen des EU-weiten Stresstests

Eine gemeinsame Fehlfunktion besteht darin, dass die Performance-Indikatoren keine genauen Informationen über die Performance liefern. Ergebnisse sind oft ungenau, nicht eindeutig oder frisiert. Banken haben in ihren Büchern Bestände, die es erlauben, das Risikopotenzial dieser Bestände einzuschätzen. Haben sie darin zu viel Spielraum, können Banken am Ende für gleichartige Bestände verschiedene Risikopotenziale annehmen, indem sie das System zu ihren Gunsten manipulieren. Die Stress-Test-Methodik hat daher Sicherheitsplafonds eingeführt, um die internen Differenzen der Bankmodelle annähernd auszugleichen. Es wurde jedoch eingewendet, dass diese Methodik einer genauen individuellen Risikobewertung im Weg steht, woraus sich Fragen ergaben nach der generellen Benutzung des Stresstests für Bewertung der Performance einer Bank. Es wurde zugegeben, dass der EBA-Stresstest methodisch in einem Korsett steckt, notwendig, um Banken zu vergleichen, aber ökonomisch ohne tieferen Sinn. Obwohl der Stresstest darauf abzielt, die Risiken und das Kapital der Banken mithilfe einer gemeinsamen Methodik gleich zu behandeln, werden auf Kosten der Genauigkeit sehr verschiedene Dinge zusammengeworfen. Der Stresstest könnte daran krankens, dass er kein sehr akkurates Bild der aktuellen Performance einer Bank wiedergibt, was zu einer ungenauen Beurteilung der Performance führen könnte. Eines der Ziele des europäischen Stresstests war es, nationale Einseitigkeiten zu beseitigen und ein europäisches Niveau gleicher Wettbewerbsbedingungen (LPF) zu schaffen. Ohne das würde der Stresstest wahrscheinlich gar nicht ernst genommen. Während die Stresstests für Banken oft frustrierend sind, wurde von allen Beteiligten zugestanden, dass insgesamt die Ergebnisse ein gerechteres Bild der Performance im Vergleich der Banken untereinander bieten. Es kann also festgestellt werden, dass der Stresstest dysfunktional ist, insofern er keine komplett exakte Einschätzung der Risiko-Performance einer Bank liefert, dass er aber Manipulationen minimiert.

## Prüfungsrituale

Obwohl die Risikoteams der Banken dem Handbuch und dem LPF gegenüber aufgeschlossen waren, waren sie teilweise frustriert über die Detailgenauigkeit und die Intensität der Untersuchung. Man sah eben nicht den Sinn der Sammlung all dieser bankspezifischen Daten. Einige beschrieben die Performance-Messungen und -informationen als symbolisch: sie stellten ein Mittel dar zur Simulation von Ordnung und Kontrolle. Ein Autor schrieb, sie dienten als Prüfungsrituale, indem sie Gewissheit vermittelten, wo sonst wenig Vertrauen herrsche. Die Tatsache, dass Banken all diese detaillierten Daten

sammeln und ausgeklügelte Modelle berechnen müssen, unterstreicht den Anspruch, dass die Überwachungsbehörden tief schürfen und gründlich arbeiten. Die paneuropäischen Stresstests symbolisierten den Übergang von einer schwachen nationalen Kontrolle zu einer unparteiischen und gründlichen Aufsicht.

Neben dem Signal der Strenge wirkte die Einfachheit des Vorgangs zu seinem Vorteil. Der Stresstest kann im Wesentlichen beschrieben werden als ein Bankrating in einer Krisensituation, das einer breiteren Öffentlichkeit leichter erklärt und vermittelt werden kann. Das erhöhte das Bewusstsein, dass die europäischen Kontrolleure die Kontrolle übernahmen, indem sie die Solidität der Banken evaluierten und klare Kapitalziele setzten. Dieser Mehrwert des Stresstests wurde von den Interviewten aufgegriffen und hervorgehoben. Der Stresstest ist eine deutlich sichtbare Performancemessung, die zum Bewusstsein und zum Vertrauen der breiteren Öffentlichkeit beiträgt, vor allem auch auf den Finanzmärkten; das ist ebenfalls ein Ziel des Stresstests. Der performative Charakter des Stresstests ist von Bedeutung. Durch die Feststellung, dass Banken gut dastehen, werden Märkte veranlasst, diese Banken auch so zu behandeln. Vertrauen in eine Organisation kann der Schlüssel sein, dass diese ihre Performance-Resultate verbessert.

Allerdings reicht es nicht aus, dass der Stresstest feststellt, eine Bank sei gesund: das muss eine glaubwürdige Feststellung sein. Anfang 2009 hatten Banken gute Ratings und scheiterten doch kurz danach. Um daraus ein glaubwürdiges Unterfangen zu machen, mussten die Regulierer sicherstellen, dass im Hintergrund die Banken wirklich „aufräumten“.

Die Banken müssen ausführliche Vorlagen mit mehr als 20.000 detaillierten Daten in verschiedenen Risikogruppen ausfüllen. Sie brauchen also in kurzer Zeit Zugang zu solchen Daten von allen Zweigstellen, müssen die Daten von verschiedenen Abteilungen harmonisieren und in kleinsten Einzelheiten erklären, wie makroökonomische Variablen ihre Aktiva beeinflussen werden. Das diente keinen rituellen oder symbolischen Zwecken, sondern veranlasste sie, ihre Selbstregulierung zu verbessern. CEOs votierten für höhere Mittelansätze, um die IT-Systeme zu verbessern und den Stresstest erfolgreich zu bestehen. Diese IT-Modernisierungen wurden dann außerhalb des Stresstests benutzt, um das tägliche Risikomanagement zu verbessern. Bessere IT-Systeme helfen den Banken, den Stresstest schneller durchzuführen, aber sie helfen auch den Banken, im alltäglichen Geschäft Probleme und Risiken schneller zu entdecken. Eine weitere Verbesserung liegt darin, dass der Stresstest Mitarbeiter verschiedener Abteilungen, auch Mitarbeiter mit Kundenkontakt und solche im Risikomanagement, näher zusammen brachte. Diese Verbindungen blieben auch

nach dem Ende des Stresstests bestehen und verbesserten so das interne Risikomanagement der Banken.

Die latente Macht, die von der EBA und der EZB ausgeübt wird, richtet sich nicht gegen die Interessen der Banken. Nach Foucault <sup>1</sup> ist Zwang nicht unbedingt schlecht. Darüber hinaus, bedeutet dieser Zwang nicht, dass die Organisation all ihrer Rechte beraubt wird und als gehirngewaschene Marionette agiert. Ganz im Gegenteil kann Machtausübung, wie von Foucault diskutiert, in eine Selbstständigkeit oder ein Verantwortungsbewusstsein der Betroffenen führen. Die Kontrolleure begnügen sich nicht damit, den Banken Wissen zu vermitteln darüber, was gesund und was riskant ist. Die Methodik der EBA beinhaltet zwar vordefinierte Kategorien, die Banken haben jedoch immer noch Spielraum für Einwendungen oder Dissens (zumindest theoretisch). Sie werden angehalten, selbst zu denken. Die Banken nutzten die Erfahrungen, die sie aus dem Stresstest gewonnen hatten.

Auf den ersten Blick scheint es, dass der Stresstest lediglich die Performance der Bank in einer Stresssituation evaluiert und entsprechend sanktioniert. Banken, die sehen, wie sich ihr Eigenkapital stark vermindert, müssen sich rekapitalisieren oder das Risiko ihres Bestands verringern. Aber das ist noch nicht alles. Der Stresstest erfasst nicht nur passiv, welche Banken gut funktionieren, sie gestalten diese Performance auch aktiv. Der Stresstest ist ein Instrument der Foucault'schen „Gouvernementalität“, das die Banken zwingt, selbstverantwortlich zu handeln und ihre Ergebnisse zu verbessern, indem sie die Performance-Indikatoren berechnen. Der Stresstest ist damit latent konstitutiver Bestandteil der Performance, die er misst.

\*

\*

\*

---

<sup>1</sup> Foucault M (2011) *The Government of Self and Others: Lectures at the Collège de France, 1982-1983*. Palgrave Macmillan.